

Д. С ПЕЛЕШЕНКО, аспирант каф. ПМ ХНУРЭ

СВЯЗЬ АТРИБУТОВ ЭЛЕМЕНТОВ РАЗНЫХ УРОВНЕЙ ИЕРАРХИЧЕСКОЙ ГРАФОВОЙ МОДЕЛИ

У статті розглянуто ієрархічна графова модель газотранспортної мережі. Така інформаційна модель дозволяє зв'язати між собою різні рівні деталізації подання предметної області та формує ієрархію моделей. Запропоновано спосіб організації зв'язку атрибутів різних рівнів ієрархічної графової моделі.

В статье рассмотрена иерархическая графовая модель газотранспортной системы. Данная информационная модель позволяет связать между собой различные уровни детализации представления предметной области и формирует иерархию моделей. Предложен способ организации связи атрибутов различных уровней иерархической графовой модели.

The hierarchical graph model of gaz transporting system was considered in this article. This information model makes possible to link different detail levels of representation of data domain and creates a hierarchy of models. Method of building link between attributes of different level of hierarchically graph model was proposed.

Введение. В настоящее время во всех сферах человеческой деятельности внедряется все больше информационных, информационно-аналитических систем, систем автоматического и автоматизированного управления. Использование таких средств позволяет повысить эффективность деятельности отдельных людей и целых предприятий. Одной из важнейших задач решаемых такими системами является задача сбора, анализа и представления большого количества разрозненных данных в виде, позволяющем человеку легко проанализировать их.

В настоящее время, не смотря на высокий уровень оснащенности объектов газотранспортной системы (ГТС) Украины комплексами автоматического и автоматизированного управления, по прежнему задачи обобщения и первичного анализа данных выполняются диспетчером. При выполнении таких задач в распоряжении диспетчера есть только средства подобные специализированным инженерным калькуляторам, многие из которых были разработаны еще в 80-х годах. Существенно новое качество диспетчерского управления может быть достигнуто путем объединения множества существующих наработок в единую систему, выполняющую первичную обработку и обобщение данных, а также их удобное представление. При построении такой системы следует учитывать иерархическую природу организационной структуры ГТС и, как следствие, необходимость различного, но не противоречивого представления данных на разных уровнях этой иерархии.

Основные элементы модели ГТС. Как описано в [1] модель ГТС представляет собой иерархию кластеров. Кластер *C* состоит из множества

узлов N_C и множества A_C упорядоченных пар узлов. Каждую такую пару узлов будем называть дугой, а порядок следования вершин направлением дуги. Дуги и узлы кластера будем называть его элементами. Множество элементов кластера C будем обозначать E_C , причем $E_C = N_C \cup A_C$.

Каждому элементу e кластера C может соответствовать определенный кластер C_e . В этом случае будем называть элемент e обобщенным элементом кластера C_e а сам кластер C_e – кластером детализации элемента e . Элемент, которому не соответствует ни один кластер, будем называть простым. Кластер C_e будем называть детализирующим кластером первого уровня по отношению к кластеру C . Кластер, содержащий только простые элементы, будем называть простым кластером. Таким образом, простой кластер представляет собой ориентированный граф [3].

Справедливо следующее соответствие между объектами ГТС и элементами иерархической графовой модели:

- кластер соответствует определенному участку ГТС, представленному с определенным уровнем детализации;
- каждый простой узел соответствует местам соединения технологических элементов;
- каждому обобщенному узлу соответствует группа технологических элементов, содержащаяся в его кластере детализации;
- каждой дуге соответствует технологический элемент или группа технологических элементов.

На самом верхнем уровне иерархической графовой модели находится кластер, состоящий из одного узла.

Как отмечено в [1] каждый элемент кластера является объектом и содержит информацию о некоторой части ГТС. Согласно [2] объект состоит из множества атрибутов (полей) описывающих его состояние и множества сообщений которые он способен обрабатывать (методов). В некоторых реализациях объектной модели, например в языке C#, в некоторых реализациях C++ и других языках понятия метода и атрибута совмещены в понятии свойство (property). Это достигается созданием для каждого атрибута двух методов, один из которых предназначен для задания значения атрибута, другой для считывания значения атрибута. Такой механизм необходим для реализации реакции объекта на изменение его атрибутов. В данной работе не рассматриваются методы объекта и реализация механизма реакции объекта на изменение его атрибутов.

На более высоких уровнях иерархии объекты описывают все большую часть реального мира, в то время как набор их атрибутов не изменяется или даже сокращается. Это происходит потому, что на верхних уровнях иерархии востребована в основном обобщенная информация. Атрибуты объектов верхних уровней иерархии, очевидно, зависят от атрибутов объектов более низких уровней. Целью данной работы является формализация связи атрибутов элементов различных уровней иерархии.

Оператор обобщения. Пусть C - некоторый кластер. Рассмотрим $e \in C$ и C' - детализирующий кластер элемента e . Объект соответствующий элементу e описывает ту же часть ГТС, что и C' , но в более обобщенной форме. Пусть R – множество значений атрибутов e_i , а R'_i – множество значений атрибутов элемента $e'_i \in C'$, тогда существует некоторый оператор

$$\Psi : \prod_i R'_i \rightarrow R, \quad (1)$$

где $\prod_i R'_i$ - декартово произведение множеств R_i

Будем называть его оператором обобщения кластера C' . Оператор Ψ осуществляет обобщение информации представленное множеством атрибутов объектов кластера C' .

Рассмотрим следующий пример. Пусть C - кластер соответствующий схеме магистрального газопровода. В таких схемах компрессорные станции (КС) обычно обозначают одним элементом не раскрывая их внутренней структуры. Пусть $e \in C$ элемент соответствующий некоторой КС, и C' детализирующий кластер элемента e . Таким образом C' задает внутреннюю структуру рассматриваемой КС. Одним из параметров КС на уровне схем магистрального газопровода будет коммерческий расход – количество газа прокачиваемого всеми агрегатами КС в единицу времени. Сама КС, в упрощенной форме, состоит из нескольких газоперекачивающих агрегатов (ГПА) соединенных параллельно. Очевидно, что коммерческий расход всей КС равен сумме коммерческих расходов всех работающих ГПА. Обозначим коммерческий расход каждого ГПА q_i , а коммерческий расход всей КС обозначим Q , тогда

$$Q = \sum q_i \quad (2)$$

выражение задающее оператор обобщения Ψ кластера C' . В приведенном примере рассмотрен только атрибут коммерческий расход, на практике оператор обобщения будет содержать несколько выражений определяющих зависимости атрибутов.

Рассмотрим другой пример пусть C' - кластер соответствующий одной из ниток магистрального газопровода. Такой кластер будет состоять из элементов соответствующих участкам трубопровода, газораспределительным станциям (ГРС), запорной арматуре и другим технологическим элементам ГТС. Пусть существует схема газопровода, на которой каждая нитка магистрального газопровода представлена одним элементом от КС до КС. Кластер соответствующий такой схеме будет содержать обобщенные дуги e_i

соответствующие ниткам. Тогда для каждой такой дуги будет существовать детализирующий кластер C'_i , соответствующий внутренней структуре каждой нитки. Каждый элемент e_i будет характеризоваться длиной l , общим геометрическим объемом v , запасом газа w и суммарным расходом на всех ГРС Q . Каждый элемент кластера C' будет также содержать параметры l_j , v_k , w_k и расход q_i . Таким образом следующая группа выражений задает оператор обобщения Ψ для C'

$$Q = \sum q_i, \quad (3)$$

где i принадлежит множеству индексов соответствующих ГРС;

$$l = \sum l_j, \quad (4)$$

где j принадлежит множеству индексов соответствующих участкам трубопровода;

$$v = \sum v_k, \quad (5)$$

$$w = \sum w_k, \quad (6)$$

где k принадлежит множеству индексов всех элементов.

Существуют более сложные операторы обобщения, но все же они могут быть представлены в виде набора выражений задающих каждый атрибут обобщенного элемента в отдельности. Другими словами, оператор обобщения можно разбить на конечное множество операторов

$$\Psi_j : \prod_k \tilde{R}_k \rightarrow A_j, \quad (7)$$

$$\tilde{R}_k \subseteq R'_k, \quad (8)$$

где $A_{i,j}$ – множество значений j -го атрибута элемента e_i (для удобства будем различать атрибуты по номерам), а R'_i – множества значений атрибутов элементов детализирующего кластера элемента e_i . Такое разделение вполне правомерно, т.к. оператор Ψ всегда можно представить в виде

$$\Psi(r) = \{\Psi_j(r)\} \quad (9)$$

где $r \in \prod R'_i$. В этом случае $\tilde{R}_k = R'_k$, а $\{k\} = \{i\}$. Однако для каждого атрибута обобщенного элемента в реальности почти всегда существует правило по которому можно определить его значение на основании значений одного-двух атрибутов элементов образующих его детализирующий кластер т.е. $|\prod \tilde{R}_k| \ll |\prod R'_i|$. Представим оператор Ψ_j в следующем виде:

$$\Psi_j(\tilde{r}) = F_j(G_1(\hat{r}_1), G_2(\hat{r}_2), \dots, G_n(\hat{r}_n)), \quad (10)$$

где $\tilde{r} \in \prod_k \tilde{R}_k$, $\hat{r}_1 \in \prod_{i'} \hat{R}_{1,i'}$, $\hat{r}_2 \in \prod_{i'} \hat{R}_{2,i'}$... $\hat{r}_n \in \prod_{i'} \hat{R}_{n,i'}$, $\hat{R}_l \subseteq \tilde{R}_l$, $\{l\} \subseteq \{k\}$,

а F и G_1, G_2, \dots, G_n – некоторые операторы.

Пусть $a_{i,j} \in A_{i,j}$ – значение j -го атрибута элемента $e_i \in C'$, а F – некоторый оператор, определенный на множестве $\Phi = \prod_k \tilde{R}_k = \prod_i \prod_j A_{i,j}$, состоящем из векторов вида:

$$(a_{1,1}, a_{1,2}, \dots, a_{1,m}, a_{2,1}, a_{2,2}, \dots, a_{2,m}, \dots, a_{p,1}, a_{p,2}, \dots, a_{p,m}), \quad (11)$$

а $I(a_{i,j}) = a_{i,j}$, тогда

$$\begin{aligned} & F(a_{1,1}, a_{1,2}, \dots, a_{1,m}, a_{2,1}, a_{2,2}, \dots, a_{2,m}, \dots, a_{p,1}, a_{p,2}, \dots, a_{p,m}) = \\ & = F(I(a_{1,1}), I(a_{1,2}), \dots, I(a_{1,m}), I(a_{2,1}), I(a_{2,2}), \dots, I(a_{2,m}), \dots, I(a_{p,1}), I(a_{p,2}), \dots, I(a_{p,m})) \end{aligned} \quad (12)$$

Если принять $F = \Psi_j$, а $G_1 = I$, $G_2 = I$, ... $G_n = I$, то выражение (12) эквивалентно выражению (10), следовательно всегда существует хотя бы один способ представить оператора Ψ_j в виде (10). Таким образом, оператор обобщения может быть представлен в виде множества операторов вида $F_j(G_1(\hat{r}_1), G_2(\hat{r}_2), \dots, G_n(\hat{r}_n))$. Такое представление оператора обобщения, как будет показано далее, более удобно в применении.

Рассмотрим пример. Пусть C – кластер, соответствующий схеме линейного участка магистрального газопровода. Линейный участок характеризуется наличием нескольких отборов и одной или нескольких подач газа. Элементы соответствующие отборам и подачам содержат атрибуты задающие расход газа. Пусть e – обобщающий элемент кластера C содержит атрибут W , задающий скорость изменения запаса газа. Тогда можно определить обобщающий оператор следующим образом:

$$W = F(G_1, G_2) = G_1(\{q_i\}) - G_2(\{q_j\}) = \sum q_i - \sum q_j, \quad (13)$$

где q_i – расходы газа j -х на подачах, а q_j – расходы газа i -х на отборах.

Обычно операторы G_i представляют собой различные агрегатные функции (суммирование, нахождение максимального значения и т.д.). Система множеств \hat{R}_i определяется на основе множества R_p значений специальных атрибутов и соответствует разбиению множества элементов кластера на подмножества M_i . В приведенном примере такими атрибутами является атрибут, определяющий является ли объект подачей газа, и атрибут определяющий является ли объект отбором газа.

Введем систему предикатов $P_i(r_p), r_p \in R_p$, которая будет задавать разбиение множества элементов кластера C на подмножества M_i :

$$M_i = \{e_{i,j} / P_i(r_{p,j}) = 1\}, \quad (14)$$

где $r_{p,j}$ - вектор значений специальных атрибутов элемента $e_{i,j}$.

В случае когда G_i - некоторая агрегатная функции множества атрибутов, действующих в каждом элементе множества M_i равны между собой. Пусть S_i - множество атрибутов каждого элемента M_i , образующее множество значений атрибутов \hat{R}_i . Следует отметить, что такое построение системы \hat{R}_i использовано в (12), где S_i состоит из одного атрибута. Таким образом каждый оператор Ψ_j может быть определен следующей структурой $\langle F_j, \{S_i\}, \{G_i\}, \{P_i\} \rangle$. А оператор Ψ - множеством таких структур. Это дает возможность на практике задавать связь атрибутов обобщающих элементов с атрибутами элементов детализирующего кластера с помощью суперпозиции некоторой функции и множества агрегатных функций над значениями атрибутов подмножеств элементов детализирующего кластера. Для разбиения на такие подмножества достаточно задать предикат на множестве значений специальных атрибутов.

Заключение. В данной работе дополнена информационная модель ГТС. Предложен способ обобщения и первичного анализа информации о ГТС основанный на иерархической структуре ГТС. В результате задания оператора Ψ_j структурой вида $\langle F_j, \{S_i\}, \{G_i\}, \{P_i\} \rangle$, на практике возможным становится создания языка позволяющего задавать оператор обобщения.

Список литературы: 1. Пелешенко Д. С. Об одном классе информационных моделей газотранспортных систем – Вестник ХГТУ – 2007.-№ 3 – С 221-228 2. Гради Буч Объектно-ориентированный анализ и проектирование с примерами приложений на C++ - издательство "Бином" 1998. - 560 стр. 3. Matthew Rapaport. Object-Oriented Data Bases: The Next Step in DBMS Evolution - Comp. Lang. - 5, N 10.- 1988.- 91-98

Поступила в редколлегию 05.11. 07